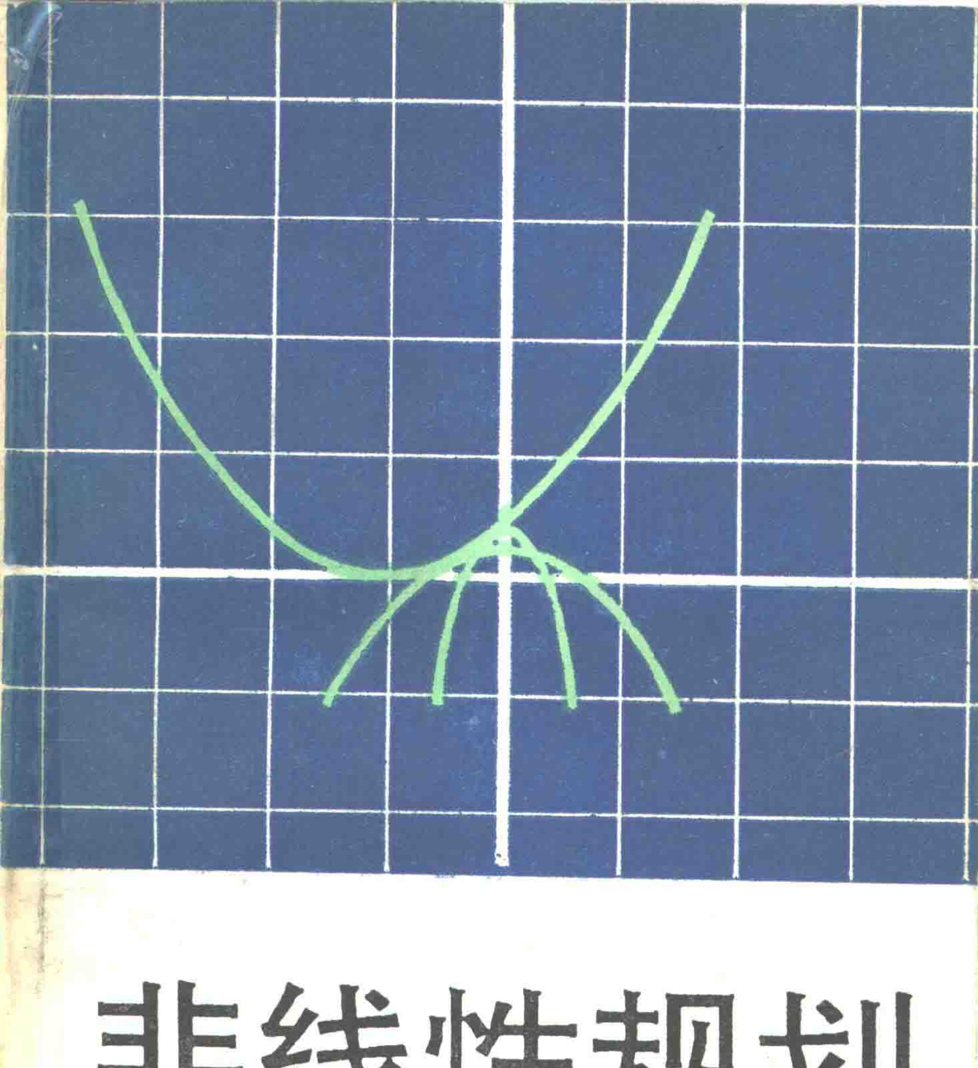

非线性规划

刘宝光 编著

北京理工大学出版社

非线性规划

刘宝光　编著

北京理工大学出版社

内 容 简 介

本书讲非线性规划的基本的、切近实用的理论结果和方法。内容的取舍遵循三个原则：1°. 突出方法。对于支撑方法的理论概念，系统条理地介绍，但不作证明；2°. 只讲近来有效地使用的方法；3°. 对方法着重讲清原理，明确算法，对收敛性等理论结果只叙述而不证明。此外，收进了应用上十分重要的实用灵敏度分析方法。讲法上有两大特点：一是注意解决同工科学生数学基础知识的衔接好的问题；一是多用几何直观，用欧氏空间的几何形象表述理论概念。

本书共九章，依次为：引论（含预备知识），凸集和凸函数，最优性条件，一维搜索法，无约束最优化方法，线性规划，线性约束最优化方法，非线性约束最优化方法，灵敏度分析。

本书是工程各专业硕士研究生学位课的教材，讲60学时。也可以供管理专业研究生，应用数学专业本科生，在职工程师等的选修课和培训班作教材用。

非线性规划

刘宝光 编著

※

北京理工大学出版社出版

新华书店北京发行所发行 各地新华书店经售

北京理工大学印刷厂印刷

※

787×1092毫米 32开本 9印张 200千字

1988年12月第一版 1988年12月第一次印刷

ISBN 7－81013－204－0/O·35

印数：1－3100 册 定价：1.80元

前　　言

本书作为教材，可供工程和管理各专业研究生、本科生修最优化课程，以及在职的工程和管理人员自学和培训使用。自 1980 年以来，我们在研究生和本科生两个层次中，开设了非线性规划或者最优化课程。曾先后编印过两稿讲义作为这些课程的教材。因为课程的大多数对象都是工程和管理专业的学生，为了适应这些读者的需求，就把教材改编成了本书。

考虑到读者都是大学三、四年级以上的，我们将微积分和线性代数的基本内容，作为本书自由引用的基础知识。几年来的教学实践表明，学生的数学知识，尚不能与课程很好地衔接。其主要原因是：(1) 学生对线性代数知识的掌握情况很不整齐；(2) 目前多元函数微分学的教学，是以偏导数为核心概念，平行地推广一元函数微分学内容，学生熟悉的只是二元函数，对于以 n 维欧氏空间、向量和矩阵等为基本概念，讨论多元函数微分学，都缺乏知识。为了解决这个衔接问题，我们在第一章中设了预备知识一节，扼要介绍这些概念。

对非线性规划内容的取舍，本书遵循三条原则。第一是突出方法。对支撑这些方法的基本理论和概念，用凸集和凸函数、最优性条件两章来加以介绍，较全面地给出可微情形的一、二阶最优性条件的理论结果（但不作严格证明，仅借助几何形象直观地加以表述）。第二是只讲近年来使用效果

好的方法。第三是在方法的讲述中，着重讲清原理，明确算法，对收敛性只给结果，不作证明。

除一般的非线性规划内容之外，考虑到灵敏度（或稳定性）分析实际上是非线性规划理论和应用的一个重要方面，本书还专辟一章，讲述较为实用的，在实际计算中可实现的灵敏度分析方法。

限于作者水平，内容和讲法的处理多有不当之处，错误也在所难免。恳请大家批评指正。

作　者

目　　录

第一章　引论

§1.1　模型问题和模型算法 …………………………（1）

1. 模型问题　2. 工程设计的优化问题　3　离散最优控制问题
4. 经营管理问题　5. 数据拟合问题　6. 模型算法
7. 算法的评价与比较

§1.2　预备知识 …………………………………（14）

1. 线性空间　2. n 维欧氏空间　3. 矩阵　4. 欧氏空间中的点集
5. 多变量函数

习　题　一 ………………………………………（25）

第二章　凸集和凸函数

§2.1　凸　集 …………………………………（29）

1. 凸组合　2. 凸集的定义和简单举例　3. 凸集的性质
4. 凸锥

§2.2　凸函数 …………………………………（36）

1. 凸函数的定义和简单例子　2. 连续性和方向可微性
3. 可微凸函数

习　题　二 ………………………………………（43）

第三章　最优性条件

§3.1　古典最优性条件 ……………………………（46）

1. 解的定义，凸规划的解　2. 无约束问题　3. Lagrange 乘子法

§3.2　Kuhn-Tucker 条件 ……………………………（53）

1. 不等式约束问题的 Kuhn-Tucker 条件　2. 最优性的几何条件
3. Kuhn-Tucker 条件定理的证明　4. 一般约束问题的 Kuhn-Tucker 条件　5. 一阶充分条件　6. 线性约束问题的一阶条件
§3.3　二阶条件 …………………………………… (70)
1. 二阶必要条件　2. 二阶充分条件
习　题　三 …………………………………… (75)

第四章　一维搜索法

§4.1　不精确一维搜索 …………………………… (79)
§4.2　0.618 法和中点法 ………………………… (82)
1. 不确定区间　2. 0.618 法　3. 中点法　4. 初始不确定区间
§4.3　Newton 法和插值法 ……………………… (88)
1. Newton 法　2. 插值法
习　题　四 …………………………………… (90)

第五章　无约束最优化方法

§5.1　收敛性概念 ………………………………… (93)
1. 算法的收敛性　2. 收敛速度　3. 二次终结性
§5.2　Newton 法及其修正 ……………………… (103)
1. 最速下降法的动机和缺点　2. Newton 法
3. 修正的 Newton 法
§5.3　变尺度法 ………………………………… (109)
1. 变尺度法模型和拟 Newton 方程　2. DFP 法
3. BFGS 法
§5.4　共轭方向法 ……………………………… (119)
1. 共轭梯度法的基本算法　2. 共轭梯度法
3. Powell 法　4. 修正的 Powell 法
习　题　五 …………………………………… (132)

第六章　线性规划

§6.1　多面体……………………………………（138）

1．极点和极方向　2．多面体的表示

§6.2　最优解和最优乘子…………………………（143）

1．标准型　2．最优性条件和对偶　3．最优解存在定理

§6.3　单纯形法……………………………………（150）

1．原理和算法　2．单纯形表　3．人工变量

习 题 六……………………………………………（160）

第七章　线性约束最优化方法

§7.1　既约梯度法…………………………………（164）

1．可行方向法模型　2．　既约梯度及方法原理

3．算法与举例

§7.2　起作用集法…………………………………（176）

1．线性等式约束　2．起作用集法原理　3．算法和举例

4．初始信息

§7.3　二次规划……………………………………（193）

1．等式约束　2．不等式约束

习 题 七……………………………………………（199）

第八章　非线性约束最优化方法

§8.1　罚函数和闸函数……………………………（202）

1．罚函数概念　2．罚函数法　3．闸函数法

§8.2　乘子法………………………………………（216）

1．乘子罚函数　2．等式约束问题的乘子法

3．向一般问题的推广

§8.3　约束变尺度法………………………………（230）

1．Lagrange-Newton 法　2．约束变尺度法

§8.4　GRG 法……………………………………………（241）

1. 原理　2. 方法

习题八……………………………………………………（253）

第九章　灵敏度分析

§9.1　基本定理…………………………………………（259）

§9.2　灵敏度分析的实现……………………………（266）

1. 无约束方法　2. 混合罚函数法　3. 乘子法

习 题 九…………………………………………………（274）

参 考 书 目…………………………………………………（277）

第一章　引　论

本章给出非线性规划问题和算法的初步概念，介绍问题的实际背景，概述欧氏空间和多变量函数理论的某些结果，为读者提供学习本书所需的数学预备知识。

§1.1　模型问题和模型算法

1. **模型问题**　非线性规划问题，是求多变量函数在特定区域上的最小问题。通常，n 元函数的自变量用一个 n 维向量 x 表示。一般的 n 元函数写作 $f(\mathrm{x})$。设 S 是 $f(\mathrm{x})$ 定义域的一个子集合。在非线性规划中，如下的书写格式

$$(fS)\qquad\begin{cases}\min & f(\mathrm{x})\\ \text{s.t.} & \mathrm{x}\in S\end{cases}$$

表示给出一个问题：求 $f(\mathrm{x})$ 在集合 S 上的最小，包括求 f 在 S 上取得最小值的点 x^*，x^* 称为该问题的**最优解**，或简称为**解**；f 在 x^* 的值 f^*，称为**最优值**。这是非线性规划的最一般的模型问题。贯穿全书，我们都用 (fS) 来记它。可以看出，构成这一模型的要素有三个，即 x，f 和 S，分别称为问题的**变量**，**目标函数**和**可行集**。可行集的每个点称为**可行解**。显然，最优解是一个特殊的可行解，是使目标函数取最优值的可行解。

对模型 (fS) 的各要素作些具体的规定，就得到较具体

一些的模型问题。例如，取 S 为整个空间，有

$$(f) \qquad \min \quad f(x)$$

如果将 S 规定为一组方程和不等式在一个大的区域 D 上的解集合，则成为

$$(fghD) \begin{cases} \min \quad f(x) \\ \text{s.t.} \quad g_i(x) \leqslant 0, \quad i=1,2,\cdots,m & (1.1) \\ \qquad h_i(x)=0, \quad i=1,2,\cdots,l & (1.2) \\ \qquad x \in D & (1.3) \end{cases}$$

模型问题 $(fghD)$ 中的不等式 (1.1)，称为**不等式约束**，方程(1.2) 称为**等式约束**，函数 g_i 和 h_i 则称为**约束函数**，式(1.1)–(1.3)合称为**约束条件**。通常将问题(f)称为**无约束问题**，将有约束条件出现的问题称为**约束问题**。

模型问题 (f)、$(fghD)$包含了非线性规划的研究主体。问题 (f) 和作为 $(fghD)$ 简单特例的

$$(fh) \begin{cases} \min \quad f(x) \\ \text{s.t.} \quad h_i(x)=0, \quad i=1,2,\cdots,l \end{cases}$$

就是古典数学分析中的极值问题。这些问题不仅是来自多元函数的研究本身，而且有很强的实际背景，因为对于“如何以最好的方式行事”这类人人感兴趣的问题，这是一种可能的数学表示。伴随着近代计算技术的飞跃发展，非线性规划提供了可以付诸实用的理论和数值方法，成为当今在工程设计、过程控制、经营管理等各个方面强有力的数学工具。关于这些领域的问题向非线性规划模型问题的归结，以举例方式

浅述于以下几段。

2. 工程设计的优化问题 一项工程设计，总是要求设计者在一定的限制条件下，通过设计来满足某几项性能要求。在选定了一种设计思想后，有一系列的设计方案可供选择。如果设计工作是充分定量化的，则不同的方案即表现为一组变量的不同取值。这组变量可称为设计变量。设计变量的取值，必须符合设计问题中给出的限制条件，也必须使所确定的设计方案满足问题所提出的各项性能要求。这些条件和要求，可以归结为由设计变量的某些函数构成的方程和不等式。这些约束条件，和设计变量，构成这种定量化设计的二要素。

如果除了通常设计问题的条件和要求外，又认定一项指标，要求所选择的设计方案，使该指标有最好的值，就成为优化设计问题。很明显，设计变量、指标和约束条件作为三

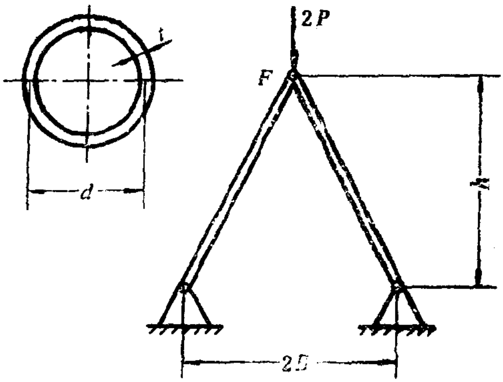

图 1.1

要素，优化设计问题就归结为（$fghD$）型的模型问题。

作为一个具体的例子，设要设计如图 1.1 所示的两杆桁架。两臂用圆钢管制成，类型已定，臂厚为 t，钢管直径不超过 D。两臂在 F 处铰接，跨度定为 $2B$。桁架高度限定不超过 H。桁架在 F 处承受载荷 $2P$，要求钢管即不超过弹性屈服限，也不发生屈曲。在满足上述所有条件和要求的前提下，要桁架的重量为最小。

这一问题中，许多条件都已固定，可以调节的因素只有钢管直径 d 和桁架高 h。这是该问题的设计变量。要求最优化的指标为重量 $w=2\pi\rho t d(B^2+h^2)^{1/2}$。问题的限制条件表现为不等式 $0\leqslant d\leqslant D$ 和 $0\leqslant h\leqslant H$。由于杆件的应力 $\sigma=P\sqrt{B^2+h^2}/(\pi t dh)$，而性能要求中要保证不超过屈服限，应力 σ 应该不超过最大允许压应力 σ_1，这产生不等式 $P\sqrt{B^2+h^2}/(\pi t dh)\leqslant\sigma_1$。不发生屈曲的要求，体现为应力 σ 不超过 $Eular$ 应力 $\sigma_e=\dfrac{(\pi^2E/8)(d^2+t^2)}{(B^2+h^2)}$。

这导出不等式 $P(B^2+h^2)^{3/2}-(\pi^2E/8)tdh(d^2+t^2)\leqslant 0$。于是，这一设计问题的数学模型为 $(fghD)$ 的特例，即

$$
\begin{cases}
\min & d(B^2+h^2)^{1/2} \\
\text{s.t.} & d-D\leqslant 0 \\
 & h-H\leqslant 0 \\
 & (B^2+h^2)^{1/2}-c_1dh\leqslant 0 \\
 & (B^2+h^2)^{3/2}-c_2dh(d^2+t^2)\leqslant 0 \\
 & d,\ h\geqslant 0
\end{cases}
$$

其中 c_1, c_2, D, H, t 等为已知常数。这一模型的含义是：

在满足这一组不等式的所有的点(d,h)中，求使函数$d(B^2+h^2)^{1/2}$取最小值的点。

3. 离散最优控制问题 设有一个分为k个阶段的过程控制问题。在第i阶段，系统用一个状态向量$\mathbf{x}_i$表示。控制向量$\mathbf{u}_i$将系统的状态从$\mathbf{x}_{i-1}$改变到$\mathbf{x}_i$，定量关系由状态方程$\mathbf{x}_i=\mathbf{x}_{i-1}+\varphi_i(\mathbf{x}_{i-1},\ \mathbf{u}_i)$给出，其中$\varphi_i$是与$\mathbf{x}_i$有相同维数的向量函数。按照状态方程给出的规律，如果初始状态$\mathbf{x}_0$给定，那么由控制向量序列$\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_k$就确定了系统的状态序列$\mathbf{x}_0,\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_k$。如图1.2所示。

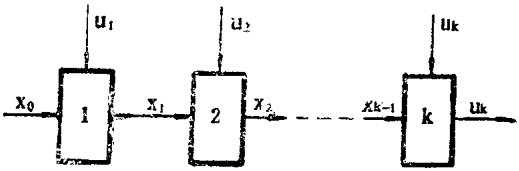

图 1.2

化学工程中的十字流多级直列串联萃取过程，就是上述一般过程的一个特例。设共有k级，示意于图1.3。图中的

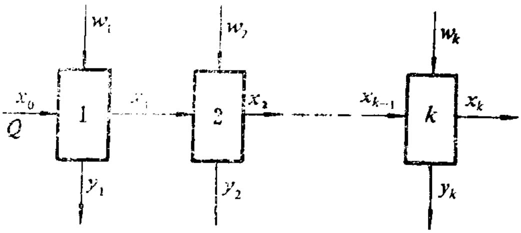

图 1.3

Q 表示进物料流量，x_0 表示料液中萃取质的浓度。在第 i 级，w_i 是溶剂流量，x_{i-1} 是进入第 i 级的料液中萃取质的浓度，y_i 是溶剂中萃取质的浓度。在这个例子中，二维的（x_i，y_i）构成第 i 级的状态向量，而控制向量仅含一个分量，即溶剂流量 w_i。状态向量是二维的，状态方程有两个：由物料平衡关系，有$Q(x_{i-1}-x_i)=w_iy_i$；由萃取的平衡关系，有 $y_i=\varphi(x_{i-1})$，其中 φ 是已知函数。

控制向量 u_i 的取值，是由操作者按计划确定的，如萃取例中的流量 w_i。不同的控制向量序列，决定不同的状态向量序列。过程的各项性能指标是状态序列和控制序列的函数。为了使某项指标具有最优值，要通过选取恰当的控制序列来实现，这就产生了最优控制问题。在这种分作 k 级最优控制问题的情形中，将控制向量序列和状态向量序列合在一起，作为问题的变量 x，要实现取最优值的那项指标是 x 的函数。状态方程组给出 x 所应满足的等式约束，除此之外，控制向量（变量 x 的一部分）还必须遵从某些限制条件。

例如，在多级萃取过程的例子中，当以总的经济收益为目标时，设 α 为萃取质单价，β 为溶剂单价，略去其他因素，将总收益表示为$\sum_{i=1}^{k}(\alpha w_iy_i-\beta w_i)$，就得到数学模型

$$
\begin{cases}
\min \quad \sum_{i=1}^{k}(\beta w_i-\alpha w_iy_i) \\
\text{s.t.} \quad Q(x_{i-1}-x_i)-w_iy_i=0, \quad i=1,2,\cdots,k \\
\qquad\quad y_i=\varphi(x_{i-1}), \quad i=1,2,\cdots,k
\end{cases}
$$

$$\begin{cases} w_i \leqslant w, & i=1,2,\cdots,k \\ w_i,\ x_i \geqslant 0, & i=1,2,\cdots,k \end{cases}$$

其中 x_0 为已知常数。它同样是属于模型问题（$fghD$），即约束问题。

4. 经营管理问题 线性规划在经营管理中的应用十分广泛，它能处理大量出现的线性问题，但在经营管理中所碰到的问题并不都是线性问题，当问题中的各有关的量不全是简单的比例关系时，就会出现非线性规划问题。这里只举一个例子，即非线性的背包问题。

经营计划决策，有时会碰到下述类型的问题。设有资金总额 w，用以投资在若干个地方设立工厂。设共有 n 个地方可供选择：在第 i 地，投资额需为 w_i，作为投资效益表征的是利润 c_i。问题是：在总投资限额 w 之内，选择哪几个地方设厂，总利润为最高？

对于第 i 地的取舍，我们用变量 x_i 表示。x_i 只取 0，1 两个值。$x_i=1$ 表示在第 i 地设厂，$x_i=0$ 表示不取第 i 地。于是，总投资为 $\sum_{i=1}^{n} w_i x_i$，总利润为 $\sum_{i=1}^{n} c_i x_i$。问题成为

$$\begin{cases} \max & \sum_{i=1}^{n} c_i x_i \\ \text{s.t.} & \sum_{i=1}^{n} w_i x_i \leqslant w \\ & x_i = 0,\ 1,\quad i=1,2,\cdots,n \end{cases}$$

这个问题常称为“背包问题”，它是因如下一个游戏性

质问题的数学模型而得名的。设有 n 个物件，第 i 件重 w_i，值 c_i。今有一背包，总载重能力为 w，问题是：选择装入哪几件，使总值为最高，这跟前边的选择建厂问题，有完全相同的数学构造。

在一个地区内，设置有n个雨量观测站。每个观测站，年降雨量这个随机变量的方差越大，该站的观测值所提供的信息量当然也越大。同时，不同站的降雨量不是独立的，从信息获取这一角度说，原来n个站的设置有一定的多余度。两站降雨量的协方差，表示两地降雨量的相关性质。所有协方差绝对值的总和越大，上述多余度则越大。依据历年来各站的观测值，可以估值出任意第 i，j 站的协方差的绝对值c_{ij}（$i \neq j$），以及第 i 站的方差 V_i。现在要考虑撤掉一些站。要求所保留的站，信息的总获得量在用方差度量的意义下，不低于 V，而使得以协方差绝对值总和度量的多余度为最小。

象前面一样引入变量 x_i，只取 0，1 两个值：取值 1 表示保留第 i 站，取 0 表示撤消第 i 站，于是，$\sum_{i=1}^{n} V_i x_i$ 表示所获信息总量，而 $\sum_{i,j=1}^{n} c_{ij} x_i x_j$（设 $c_{ii}=0$）则为多余度。

问题成为

$$
\begin{cases}
\min & \sum_{i,j=1}^{n} c_{ij} x_i x_j \\
\text{s.t.} & \sum_{i=1}^{n} V_i x_i \geq V \\
 & x_i = 0,\ 1,\quad i = 1, 2, \cdots, n
\end{cases}
$$